CCA-500 Q&As

Cloudera Certified Administrator for Apache Hadoop (CCAH)

Pass Cloudera CCA-500 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.passapply.com/cca-500.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Cloudera Official Exam Center

QUESTION 1

You observed that the number of spilled records from Map tasks far exceeds the number of map output records. Your child heap size is 1GB and your io.sort.mb value is set to 1000MB. How would you tune your io.sort.mb value to achieve maximum memory to disk I/O ratio?

A. For a 1GB child heap size an io.sort.mb of 128 MB will always maximize memory to disk I/O

B. Increase the io.sort.mb to 1GB

C. Decrease the io.sort.mb value to 0

D. Tune the io.sort.mb value until you observe that the number of spilled records equals (or is as close to equals) the number of map output records.

Correct Answer: D

QUESTION 2

Identify two features/issues that YARN is designated to address: (Choose two)

A. Standardize on a single MapReduce API

B. Single point of failure in the NameNode

C. Reduce complexity of the MapReduce APIs

D. Resource pressure on the JobTracker

E. Ability to run framework other than MapReduce, such as MPI

F. HDFS latency

Correct Answer: DE

QUESTION 3

You have recently converted your Hadoop cluster from a MapReduce 1 (MRv1) architecture to MapReduce 2 (MRv2) on YARN architecture. Your developers are accustomed to specifying map and reduce tasks (resource allocation) tasks when they run jobs: A developer wants to know how specify to reduce tasks when a specific job runs. Which method should you tell that developers to implement?

A. MapReduce version 2 (MRv2) on YARN abstracts resource allocation away from the idea of "tasks" into memory and virtual cores, thus eliminating the need for a developer to specify the number of reduce tasks, and indeed preventing the developer from specifying the number of reduce tasks.

B. In YARN, resource allocations is a function of megabytes of memory in multiples of 1024mb. Thus, they should specify the amount of memory resource they need by executing D mapreducereduces.memory-mb-2048

C. In YARN, the ApplicationMaster is responsible for requesting the resource required for a specific launch. Thus, executing D yarn.applicationmaster.reduce.tasks=2 will specify that the ApplicationMaster launch two task contains on

the worker nodes.

D. Developers specify reduce tasks in the exact same way for both MapReduce version 1 (MRv1) and MapReduce version 2 (MRv2) on YARN. Thus, executing D mapreduce.job.reduces-2 will specify reduce tasks.

E. In YARN, resource allocation is function of virtual cores specified by the ApplicationManager making requests to the NodeManager where a reduce task is handeled by a single container (and thus a single virtual core). Thus, the developer needs to specify the number of virtual cores to the NodeManager by executing p yarn.nodemanager.cpu-vcores=2

Correct Answer: D

QUESTION 4

Table schemas in Hive are:

A. Stored as metadata on the NameNode

B. Stored along with the data in HDFS

C. Stored in the Metadata

D. Stored in ZooKeeper

Correct Answer: B

QUESTION 5

For each YARN job, the Hadoop framework generates task log file. Where are Hadoop task log files stored?

A. Cached by the NodeManager managing the job containers, then written to a log directory on the NameNode

B. Cached in the YARN container running the task, then copied into HDFS on job completion

C. In HDFS, in the directory of the user who generates the job

D. On the local disk of the slave mode running the task

Correct Answer: D